Cell PRESS
Structure
Short Article

Structure of the DNA-Binding and RNA-Polymerase-Binding Region of Transcription Antitermination Factor λQ

Sergey M. Vorobiev,[1,6] Yocheved Gensler,[2,6] Hanif Vahedian-Movahed,[3] Jayaraman Seetharaman,[1] Min Su,[1] Janet Y. Huang,[4,5] Rong Xiao,[4,5] Gregory Kornhaber,[4,5] Gaetano T. Montelione,[4,5] Liang Tong,[1] Richard H. Ebright,[3,*] and Bryce E. Nickels[2,*]

[1]Department of Biological Sciences, Northeast Structural Genomics Consortium, Columbia University, New York, NY 10027, USA
[2]Department of Genetics and Waksman Institute, Rutgers University, Piscataway, NJ 08854, USA
[3]Department of Chemistry and Waksman Institute, Rutgers University, Piscataway, NJ 08854, USA
[4]Center for Advanced Biotechnology and Medicine, Rutgers University, Robert Wood Johnson Medical School, Rutgers University, Piscataway, NJ 08854, USA
[5]Northeast Structural Genomics Consortium, Rutgers University, Piscataway, NJ 08854, USA
[6]These authors contributed equally to this work
*Correspondence: ebright@waksman.rutgers.edu (R.H.E.), bnickels@waksman.rutgers.edu (B.E.N.)
http://dx.doi.org/10.1016/j.str.2013.12.010

SUMMARY

The bacteriophage λ Q protein is a transcription antitermination factor that controls expression of the phage late genes as a stable component of the transcription elongation complex. To join the elongation complex, λQ binds a specific DNA sequence element and interacts with RNA polymerase that is paused during early elongation. λQ binds to the paused early-elongation complex through interactions between λQ and two regions of RNA polymerase: region 4 of the σ^{70} subunit and the flap region of the β subunit. We present the 2.1 Å resolution crystal structure of a portion of λQ containing determinants for interaction with DNA, interaction with region 4 of σ^{70}, and interaction with the β flap. The structure provides a framework for interpreting prior genetic and biochemical analysis and sets the stage for future structural studies to elucidate the mechanism by which λQ alters the functional properties of the transcription elongation complex.

INTRODUCTION

Transcription is carried out by multisubunit RNA polymerases (RNAPs) that are conserved in all organisms. Transcription consists of three phases: initiation, elongation, and termination. Bacterial RNAP consists of a core enzyme (subunit composition $\alpha_2\beta\beta'\omega$) that contains determinants for random, nonspecific initiation and for elongation. To carry out promoter-specific initiation, bacterial RNAP core must associate with the initiation factor σ to form RNAP holoenzyme (subunit composition $\alpha_2\beta\beta'\omega\sigma$). σ contains determinants for sequence-specific interaction with promoter DNA. The principal bacterial σ factor, σ^{70} in *Escherichia coli*, contains five conserved regions: σR1, σR2, σR3, the σR3/σR4 linker, and σR4. σR2 and σR4 are structured, independently folded domains that mediate sequence-specific interactions with the promoter −10 element and the promoter −35 element, respectively. The interface between the core subunits and σ in RNAP holoenzyme is extensive and includes contacts that facilitate interaction between σ and the promoter −10 and −35 elements. In particular, interaction between σR2 and a domain of β′ known as the "clamp helices" is required for σR2 to bind the −10 element (Young et al., 2001). Furthermore, interaction between σR4 and the tip of a region of β known as the "flap" is required for σR4 to bind a −35 element separated by ~17 bp from the −10 element (Kuznedelov et al., 2002).

A large body of work has provided a detailed picture of the mechanisms of activators and repressors that affect transcription initiation. In contrast, relatively less is understood regarding mechanisms of factors that regulate the postinitiation stages of transcription. The bacteriophage λ Q antiterminator protein (λQ) provides a well-established model for achieving a mechanistic understanding of the postinitiation regulation of RNAP function (Roberts et al., 1998). λQ is an operon-specific elongation factor required for expression of the phage late genes, which are contained in an ~26 kB operon under the control of the λ late promoter, $\lambda P_{R'}$. λQ-mediated regulation of late gene transcription involves formation of a highly stable λQ-containing elongation complex (Deighan and Hochschild, 2007; Shankar et al., 2007; Yarnell and Roberts, 1999) that exhibits increased processivity and is resistant to terminators within the late gene operon (Roberts et al., 1998) (Figure 1).

The process that results in the formation of the λQ-containing elongation complex, "λQ-loading," occurs during early elongation and requires a network of protein-protein and protein-DNA interactions (Figure 1). λQ-loading involves three *cis*-acting DNA sequence elements embedded within the $\lambda P_{R'}$ promoter: (1) a σ-dependent pause element, (2) a λQ binding element (QBE), and (3) a TTGACT motif. The pause-inducing element is located in the initial transcribed region of $\lambda P_{R'}$, resembles a promoter −10 element, and mediates pausing of RNAP holoenzyme in transcription complexes carrying a 16 nt or 17 nt RNA

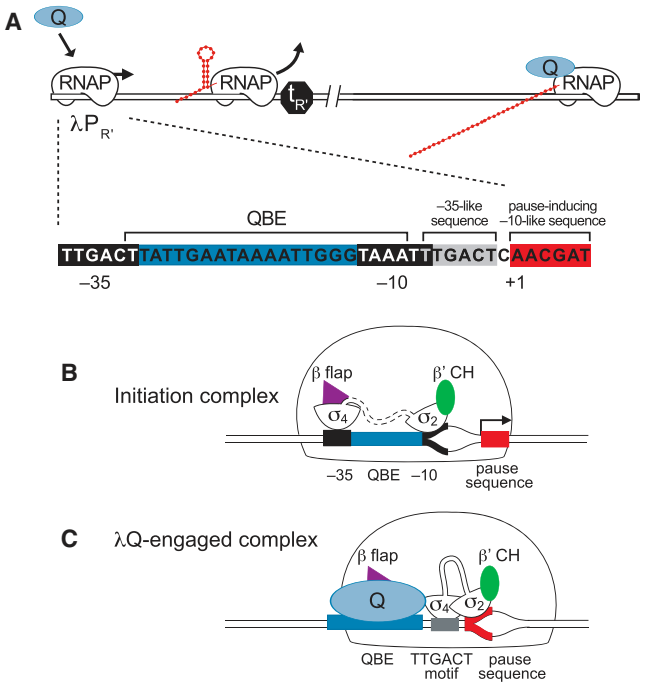

Figure 1. λQ Regulates Gene Expression from $\lambda P_{R'}$

(A) λQ (depicted as a blue oval) enables RNAP that initiates transcription from the phage late promoter, $\lambda P_{R'}$, to read through terminator $t_{R'}$ located upstream of the phage late genes. Highlighted are the $\lambda P_{R'}$ −35 and −10 elements, the QBE, the TTGACT motif (−35-like element; gray), and the pause-inducing −10-like sequence (red).

(B) Initiation complex at $\lambda P_{R'}$. σR4 (σ_4) is bound to the β flap (purple triangle). σR2 (σ_2) is bound to the β′ clamp helices (β′ CH; green oval).

(C) λQ-engaged complex at $\lambda P_{R'}$. λQ interacts with the β flap tip and stabilizes the binding of σR4 to the TTGACT motif. RNAP holoenzyme adopts a conformation with σR4 displaced from the β flap tip.

through interaction with σR2 (Ring et al., 1996). λQ engages the paused RNAP holoenzyme when bound to the QBE, which is located 8 bp upstream of the σ-dependent pause element. λQ's engagement with the paused complex involves at least two protein-protein interactions between λQ and RNAP. One of these interactions occurs between λQ and σR4 (Nickels et al., 2002b). The interaction between λQ and σR4 stabilizes the binding of σR4 to the TTGACT motif, which resembles a promoter −35 element though it is located just 1 bp upstream of the pause-inducing −10-like element (in contrast to a promoter −35 element, which is located ~17 bp upstream of a promoter −10 element) (Nickels et al., 2002b). A second protein-protein interaction occurs between λQ and the β flap tip (Deighan et al., 2008)—i.e., the same region of RNAP core that interacts with σR4 in RNAP holoenzyme. The β flap constitutes a large portion of the RNA exit channel through which the nascent RNA is extruded during transcription elongation, and the flap tip is located at the exterior opening of this channel (Vassylyev et al., 2007). Biochemical data indicate that λQ, along with the cofactor NusA, alters the sensitivity of the elongation complex to termination signals by forming a "protective barrier" around the nascent RNA transcript as it emerges from the RNA exit channel (Shankar et al., 2007). λQ's ability to form this protective barrier is likely facilitated, in part, by interaction with the β flap tip.

Genetic and biochemical studies have elucidated many aspects of λQ-dependent regulation of late gene transcription. Nevertheless, a precise mechanistic understanding of the process remains to be defined, in part, because of a lack of structural information for λQ. Here we present the structure of a portion of λQ that comprises determinants for sequence-specific DNA binding, protein-protein interaction with σR4, and protein-protein interaction with the β flap tip. The structure provides a framework for interpreting prior genetic and biochemical analyses of λQ's interactions with DNA and RNAP and represents a critical step toward obtaining a structural picture of the process by which λQ regulates late gene transcription.

RESULTS

Crystal Structure of $\lambda Q^{39–207}$

In silico predictions of disordered regions (Figure S1 available online) and small-scale tests of protein expression and solubility led to selection of a λQ fragment that includes residues 39–207, $\lambda Q^{39–207}$, for structural studies. The crystal structure of $\lambda Q^{39–207}$ was determined at 2.1 Å resolution using the selenomethionyl single-wavelength anomalous diffraction (SAD) method. The crystal contains two copies of λQ, chains A and B, in the asymmetric unit, which have similar conformations (root-mean-square [rms] distance of 0.86 Å among equivalent Cα atoms). Residues 142–157 show recognizable conformational differences between the A and B chains, and the rms distance is 0.7 Å excluding this region. Electron density was interpretable for residues 62–206 of λQ for the A chain and residues 62–204 for the B chain. The first 23 residues of the recombinant protein (residues 39–61) are not visible in the electron density map. The final R factor of the atomic model is 21.4%, and free R is 24.9% (Table 1).

The protein fragment ordered in the crystal structure has an elongated shape, with dimensions of 53 Å × 35 Å × 22 Å (Figure 2). The structure contains a tight bundle of four potentially mobile helices (α1–α2 and α4–α5). A long, arm-like structure (residues 114–152) is inserted between helices α2 and α4, which contains a short, two-stranded antiparallel β sheet and a short helix (α3). There is also a zinc ion bound at the base of the arm, coordinated by four cysteine residues (Cys118, Cys121, Cys144, and Cys147), which may be important for the structural integrity and/or flexibility of this arm. Serine and histidine substitutions at each of these four cysteine residues disrupt λQ's ability to function as an antiterminator (Guo and Roberts, 2004). Analysis of the electrostatic surface potential of λQ (Figure 2C) reveals a negatively charged patch containing the C terminus of α5 and a positively charged patch containing the C terminus of α1, the N terminus of α5, and the N terminus of α4.

A Dali search (Holm and Sander, 1995) for structurally similar proteins was performed after removing the arm-like structure (residues 114–152) between helices α2 and α4. Acidianus Filamentous Virus 1 coat protein (Protein Data Bank [PDB] ID 3FBZ; root-mean-square deviation [rmsd] 3.2 Å; Goulet et al., 2009) showed the highest structural similarity to λQ (Dali Z score 7.1; number of aligned residues 82; number of residues in target 123). However, the structural similarity is unaccompanied by significant sequence similarity and is restricted to the four-helix bundle. Moreover, three of four helices in the bundle are different

Table 1. Data Collection and Refinement Statistics	
Data Collection	
Space group	$P3_2$
X-ray wavelength (Å)	0.97913
Unit-cell parameters (Å)	$a = b = 54.5$, $c = 109.8$
Resolution (Å)	43.4-2.1
No. of unique reflections	42,458 (4,027)
Completeness	99.6 (96.8)
Redundancy	3.2 (2.1)
$I/\sigma(I)$	16.1 (1.3)
R_{merge}	0.082 (0.666)
Refinement	
Resolution range (Å)	43-2.1
R_{work}/R_{free} [a]	0.214/0.249
No. of subunits in ASU	2
No. of protein non-H atoms	2372
No. of water molecules	162
No. of heterogen atoms	8
Mean main chain B-factor (Å^2)	35.6
Mean overall B-factor (Å^2)	39
Rmsd in bond lengths (Å)	0.007
Rmsd in bond angles (°)	1.4
Ramachandran Plot	
Most favored regions (%)	95.7
Additional allowed regions (%)	4.3
Generously allowed regions (%)	0

Values in parentheses are for the highest resolution shell (2.10–2.18 Å). Reflections with F ≥ 1.05 σF were included in calculating R-factors. Data collection and refinement statistics come from *SCALEPACK* (Otwinowski and Minor, 1997) and *PHENIX* (Adams et al., 2002), respectively.
[a] *R*free is calculated for a randomly chosen set (5%) of reflections.

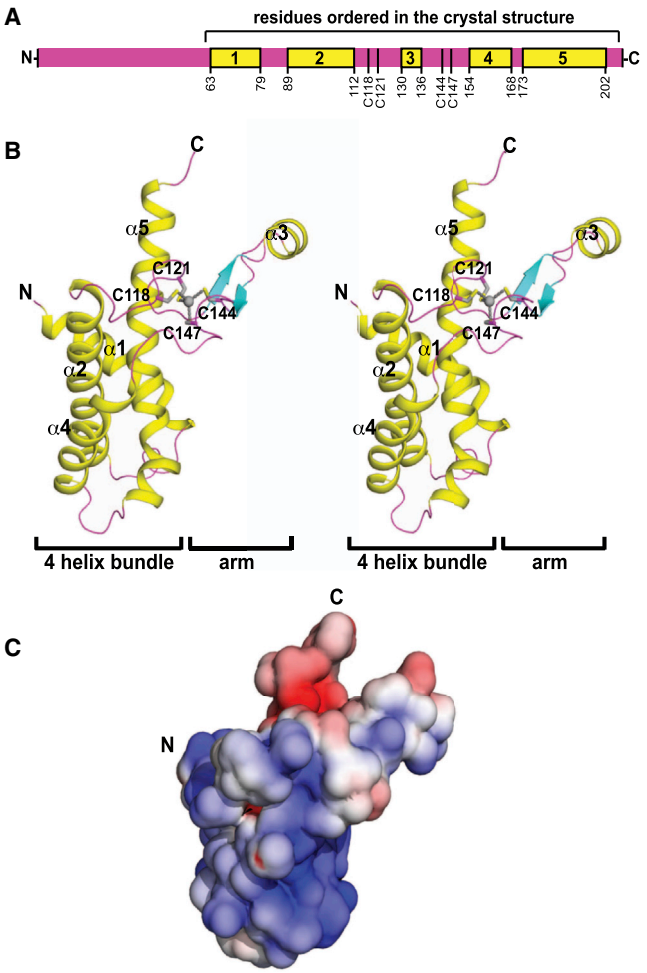

Figure 2. Crystal Structure of λQ$^{39–207}$

(A) Linear map of λQ^{FL}. Indicated is the portion of λQ ordered in the crystal structure, locations of α helices 1–5 (yellow), and location of the cysteine residues (118, 121, 144, and 147) of the metal binding site.

(B) Stereoview representation of the crystal structure of λQ prepared with PyMOL Version 1.6.0.4. β strands are colored cyan, α helices yellow, and the loops magenta. The Zn ion is shown as a sphere and colored gray. Cys residues of the metal binding site are shown as sticks (carbon in gray and sulfur in yellow).

(C) Electrostatic surface potential (Baker et al., 2001) of λQ oriented as in the left panel of (B), with fully saturated colors representing negative (red) and positive (blue) potentials of ± 4 kT.

See also Figure S1.

in length in the two proteins, and connectors between the helices are different as well.

DNA-Binding Properties of λQ$^{39–207}$ and λQ$^{62–207}$

We wished to establish whether the protein fragment used for crystallization, λQ$^{39–207}$, and a protein fragment lacking the N-terminal residues of λQ$^{39–207}$ that are not ordered in the crystal structure, λQ$^{62–207}$, retained biochemical properties of full-length λQ, λQ^{FL}. First, we performed electrophoretic mobility shift assays, which showed that λQ$^{39–207}$ retains full or nearly full sequence-specific DNA-binding activity in vitro (apparent K_d = 11 nM for λQ$^{39–207}$; apparent K_d = 4 nM for λQ^{FL}; Figure S2). Next, we tested whether or not λQ$^{39–207}$ and λQ$^{62–207}$ retained the ability to bind the QBE in vivo using a bacterial one-hybrid assay (Nickels et al., 2002a) (Figure 3A). In this assay, contact between λQ that has been fused to the α subunit of RNAP activates transcription of a test promoter (fused to a *lacZ* reporter gene) that contains a QBE upstream of the core promoter elements (Figure 3A). We introduced a plasmid that directed the synthesis of α not fused to λQ, α-λQ^{FL}, α-λQ$^{39–207}$, or α-λQ$^{62–207}$ into cells carrying a test promoter with an upstream QBE (pQBE−64) or cells carrying a test promoter without an upstream QBE (pΔQBE) (Figure 3A). Production of each of the three α-λQ fusions led to a significant increase (~3- to 7-fold) in transcription (as measured by β-galactosidase assay) from pQBE−64 and did not increase transcription from pΔQBE (Figure 3B). We conclude that λQ$^{39–207}$ and λQ$^{62–207}$ retain the determinants for sequence-specific protein-DNA interaction with the QBE.

The experiments here do not define the stoichiometry of the λQ-QBE interaction. Given the length of the λQ fragment ordered in the crystal structure (~40 Å) and the length of the QBE (~22 bp; ~76 Å), it seems probable that the stoichiometry is 2:1. Our results do not identify a candidate region that mediates protein-protein interactions between pairs of λQ monomers

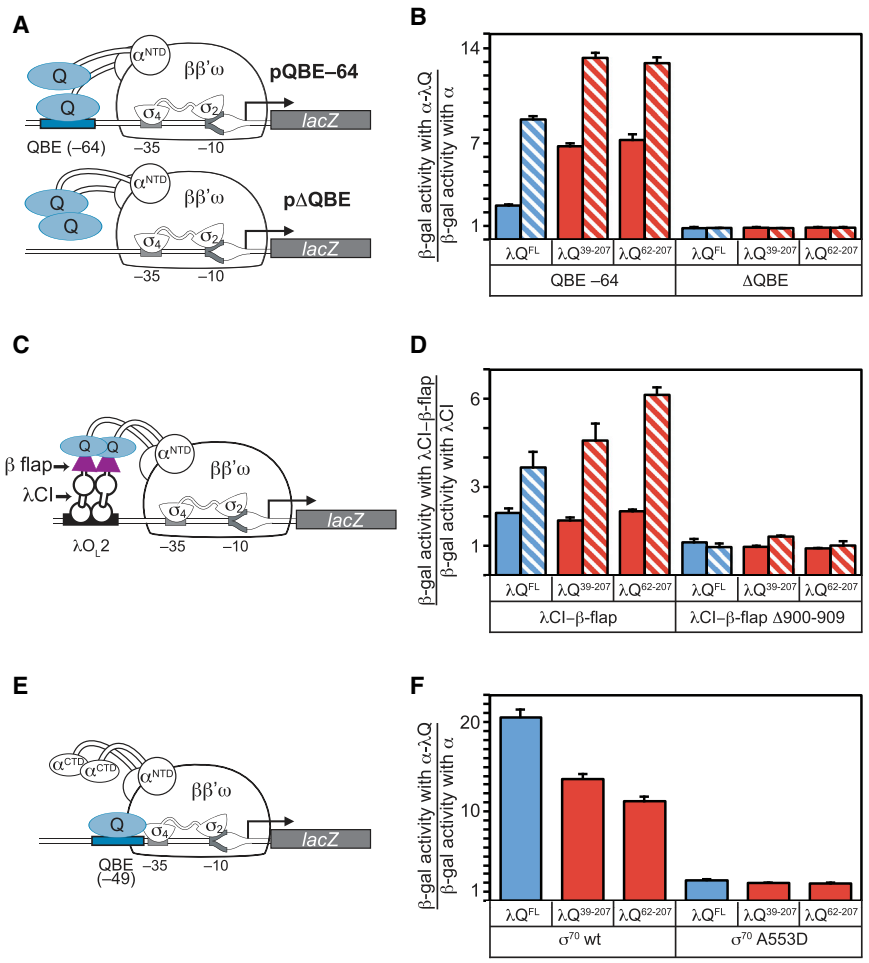

Figure 3. $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ Retain Biochemical Properties of λQ^{FL}

(A) Test promoters pQBE−64 and pΔQBE.

(B) Graph shows the ratio of β-galactosidase activity in the presence of the indicated α-λQ fusion to β-galactosidase activity in the absence of Q (α). Hatched bars indicate α-λQ contains substitution E134K. Cells were grown using medium containing 5 μM IPTG. See also Figure S2.

(C) Bacterial two-hybrid assay.

(D) Graph shows the ratio of β-galactosidase activity observed in assays performed with cells containing the indicated α-λQ fusion and the indicated λCI-β-flap fusion compared to the activity observed in assays performed with cells containing the indicated α-λQ fusion and λCI alone. Hatched bars indicate α-λQ contains substitutions T101I, A160V, and T165A. Cells were grown using medium containing 100 μM IPTG. See also Figure S3.

(E) Test promoter pQBE−49.

(F) Graph shows the ratio of β-galactosidase activity in the presence of Q (α-λQ) to β-galactosidase activity in the absence of Q (α). Cells carried either wild-type σ^{70} or σ^{70} with substitution A553D. Cells were grown using medium containing 100 μM IPTG.

Graphs represent the values and SEM obtained from quadruplicate measurements.

bound to the QBE, but that such interactions occur cannot be excluded.

Prior work has identified an amino acid substitution (E134K) that increases the binding of full-length λQ to the QBE (Guo and Roberts, 2004). Introduction of the E134K substitution into α-λQ^{FL} led to an ~3.5-fold increase in transcription compared to that observed with the wild-type α-λQ^{FL} fusion, and its introduction into the α-$\lambda Q^{39–207}$ or α-$\lambda Q^{62–207}$ fusion led to an ~2-fold increase (Figure 3B). We conclude that the E134K substitution has a similar effect on DNA binding in the context of λQ^{FL}, $\lambda Q^{39–207}$, and $\lambda Q^{62–207}$.

RNAP Binding Properties of $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$

λQ-loading also involves protein-protein interaction between QBE-bound λQ and the β flap tip. We next determined whether or not $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ retained determinants for interaction with the β flap tip. To do this, we used a bacterial two-hybrid assay previously used to detect the interaction between λQ^{FL} and the β flap tip and to isolate amino acid substitutions in λQ^{FL} that strengthen the λQ-β-flap-tip interaction (Deighan et al., 2008). In this assay, contact between λQ fused to the α subunit of RNAP and β residues 831–1,057 fused to the CI protein of bacteriophage λ (λCI) activates transcription from a test promoter that contains an upstream recognition site for λCI (Figure 3C). Although residues 831–1,057 of β comprise the entire β flap region, prior studies establish that the interaction between λQ and $\beta^{831–1,057}$ detected in this two-hybrid assay involves contact between λQ and β residues 900–909 within the β flap tip (Deighan et al., 2008).

We introduced a plasmid that directs the synthesis of α-λQ^{FL}, α-$\lambda Q^{39–207}$, or α-$\lambda Q^{62–207}$ into reporter strain cells containing either the λCI-β-flap fusion or λCI alone. The amount of transcription from the test promoter was ~2-fold higher in cells containing α-λQ^{FL}, α-$\lambda Q^{39–207}$, or α-$\lambda Q^{62–207}$ along with the λCI-β-flap fusion compared with cells containing α-λQ^{FL}, α-$\lambda Q^{39–207}$, or α-$\lambda Q^{62–207}$ along with λCI alone (Figure 3D). Furthermore, the stimulatory effect of α-λQ^{FL}, α-$\lambda Q^{39–207}$, or α-$\lambda Q^{62–207}$ was eliminated when residues 900–909 of the β moiety of the λCI-β-flap fusion were deleted (Figure 3D). These findings suggest that $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ interact with the β flap tip.

Next, we determined the effect of amino acid substitutions that strengthen the interaction between λQ^{FL} and the β flap tip (T101I, A160V, and T165A) (Deighan et al., 2008) and found that these substitutions have a similar effect on the interaction with the β flap tip in the context of λQ^{FL}, $\lambda Q^{39–207}$, and $\lambda Q^{62–207}$ (Figure 3D). We conclude that $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ contain determinants for protein-protein interaction with the β flap tip.

σ Binding Properties of $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$

λQ-loading involves protein-protein interaction between QBE-bound λQ and σR4 bound to the TTGACT motif. To test whether or not $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ retained determinants for interaction with σR4, we utilized a test promoter (pQBE−49) in which

the QBE is centered at position −49 and the TTGACT motif serves as the promoter −35 element (Nickels et al., 2002b) (Figure 3E). We have previously shown that λQ stimulates transcription from pQBE−49 and, furthermore, that this stimulatory effect requires λQ to bind to the QBE and contact σR4 (Nickels et al., 2002b).

We chose to use the α-λQ^{FL}, α-$\lambda Q^{39–207}$, and α-$\lambda Q^{62–207}$ fusions for these assays because $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ (not fused to α) were significantly less stable than λQ^{FL} (not fused to α), whereas α-λQ^{FL}, α-$\lambda Q^{39–207}$, and α-$\lambda Q^{62–207}$ exhibited similar stabilities (Figure S3). Production of the α-$\lambda Q^{39–207}$ fusion led to an ~13-fold increase in transcription from pQBE−49 compared to the amount of transcription observed upon overproduction of α alone, whereas production of the α-$\lambda Q^{62–207}$ fusion led to an ~11-fold increase (Figure 3F). The stimulatory effect of α-$\lambda Q^{39–207}$ or α-$\lambda Q^{62–207}$ on transcription from pQBE−49 was comparable to that observed with an α-λQ^{FL} fusion (~20-fold; Figure 3F). The increase in transcription from pQBE−49 observed in the presence of α-λQ^{FL}, α-$\lambda Q^{39–207}$, and α-$\lambda Q^{62–207}$ was significantly reduced in reporter strain cells that carried a mutation in σR4, A553D, which disrupts the interaction between λQ and σR4 (Figure 3F). We conclude that $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ contain determinants for protein-protein interaction with σR4.

Antitermination Properties of $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$

We next determined whether or not $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ retained the ability to function as antiterminators for transcripts initiating from $\lambda P_{R'}$. To do this, we used a reporter strain carrying a $\lambda P_{R'}$-*lacZ* fusion that measures the ability of λQ to function as an antiterminator (Figure S4A). We found that whereas ectopic production of α-λQ^{FL} led to an ~4-fold increase in *lacZ* expression, ectopic production of either α-$\lambda Q^{39–207}$ or α-$\lambda Q^{62–207}$ only slightly increased *lacZ* expression (~1.5-fold) (Figure S4A). We conclude that both $\lambda Q^{39–207}$ and $\lambda Q^{62–207}$ are missing at least some of the determinants required for them to function as antiterminators for transcripts initiating from $\lambda P_{R'}$. These findings are consistent with prior work indicating that substitutions at several positions within the first 61 amino acids of λQ (V6E, G39R, A50P, C53R, H56R, and L58P) reduce λQ's ability to antiterminate transcription while not affecting λQ's ability to bind the QBE (Deighan and Hochschild, 2007; Guo, 1999; Guo and Roberts, 2004).

DISCUSSION

Here we describe the structure of 145 C-terminal residues of the bacteriophage λ Q antiterminator protein. We establish that the protein fragment ordered in the crystal structure retains wild-type activity with respect to DNA binding, protein-protein interaction with σR4, and protein-protein interaction with the β flap tip. The structure provides a framework for interpreting the extensive prior genetic and biochemical analysis of λQ's interactions with DNA and RNA polymerase.

Mapping Residues Important for DNA Binding onto the Structure of λQ

Gou and Roberts defined a 27-residue region of λQ (amino acids 155 to 181) that determines DNA-binding specificity (Guo, 1999; Guo and Roberts, 2004). This 27-residue specificity-determining segment was contained within a 36-residue segment (amino acids 155 to 190) predicted to fold as two α helices. Alanine scanning of the second helix of the predicted helix-turn helix (HTH) identified substitutions that resulted in defects in DNA binding at residues L171, Q173, T175, W176, S177, R178, V180, K181, L183, Y184, D185, L187, and V188 (Guo and Roberts, 2004).

The structure of λQ shows that residues 155 to 168 of the 27-residue specificity-determining segment fold as a short first α helix (α4), whereas residues 172 to 181 are part of a long second α helix (α5) arranged in a manner similar, but not identical, to a canonical HTH DNA-binding motif (Figure 4A). The majority of the alanine substitutions that disrupt DNA binding map onto the second α helix of the two α helices (DNA interaction region, or DIR; red residues in Figures 4B and 4C). We propose that the second α helix constitutes the primary DNA-binding determinant for λQ-QBE interaction, and we speculate that the second α helix may interact with the DNA major groove of a segment of the QBE in a manner similar to a canonical HTH DNA-binding motif.

Guo and Roberts also identified substitutions that increased the DNA-binding affinity of λQ: at residues E134, V189, and H192 (Guo, 1999; Guo and Roberts, 2004). Two of these residues, V189 and H192, are located in or near the primary DNA-binding determinant in the second helix of the HTH, adjacent to positions where alanine substitutions affect DNA binding (pink residues in Figures 4B and 4C). The third residue, E134, is located far, at least 20 Å, from the primary DNA-binding determinant (magenta residue in Figures 4B and 4C). The distance between E134 and the primary DNA-binding determinant appears too large to enable E134 to interact with the same set of nucleotides within the QBE as the primary DNA-binding determinant. It is possible that a significant conformational change in λQ may occur upon DNA binding that localizes E134 closer to the primary DNA-binding determinant. Alternatively, it is possible that E134 interacts with a second set of nucleotides 4 to 8 bp away from the nucleotides recognized by the primary DNA-binding determinant. As will be discussed below, E134 is part of the region of λQ that interacts with σR4 when σR4 is bound to the TTGACT motif. The TTGACT motif is centered 5 bp downstream of the QBE (Figure 1A). The increased DNA-binding affinity of the E134K λQ mutant may reflect the creation of a new favorable electrostatic interaction between a positively charged lysine at position 134 and a negatively charged DNA phosphate in or near the σR4-TTGACT motif interface.

Mapping Residues Important for Interaction with RNAP onto the Structure of λQ

Deighan et al. identified substitutions at three residues of λQ, T101, A160, and T165, that strengthen the interaction with the β flap tip (Deighan et al., 2008). Although these positions are separated in the primary structure, in the tertiary structure these three positions clearly cluster (green residues in Figures 4B and 4C). Two of the residues, A160 and T165, are located in the first α helix of the HTH implicated in λQ-QBE recognition, and the third residue, T101, is located in an α helix that packs adjacent and antiparallel to the first α helix of the HTH. We refer to these three residues as the flap interaction region (FIR; green residues in Figures 4B and 4C). The surface defined by the DIR (red residues in

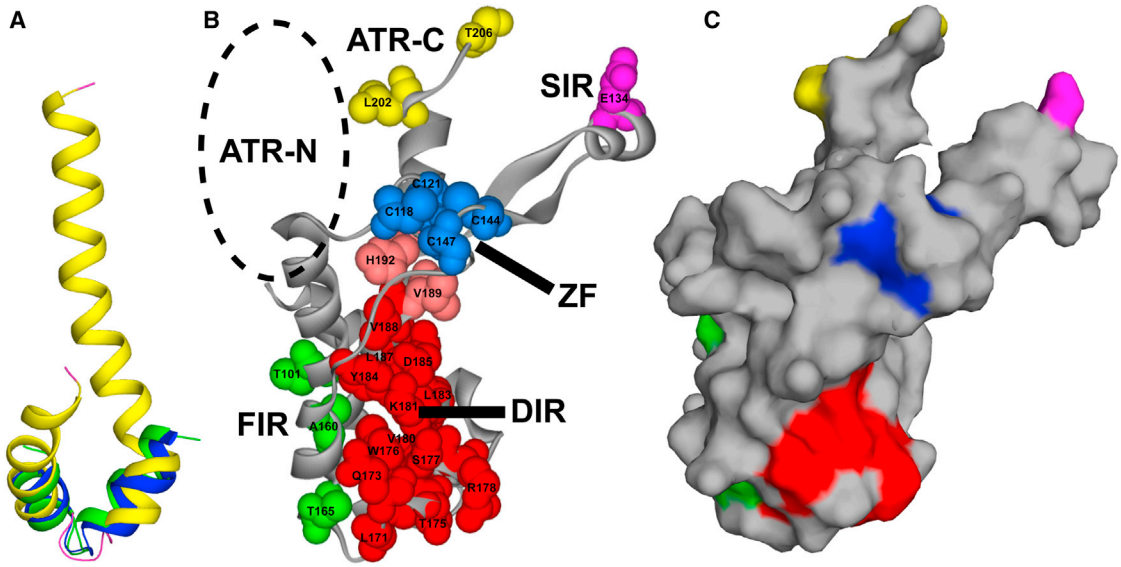

Figure 4. Functional Modules of λQ

(A) Least-squares superposition of residues 160–180 of λQ (yellow and magenta) with the HTH DNA-binding motifs of λ Cro (blue; rmsd 2.23 Å; residues 15–34 from PDB 5CRO) and λ CI (green; rmsd 2.8 Å; residues 32–52 from PDB 1LRD).

(B and C) Highlighted are the DNA interaction region (DIR; red and pink residues), RNAP flap interaction region (FIR; green residues), σ interaction region (SIR; magenta residue), C-terminal antitermination region (ATR-C; yellow residues), and Cys4 zinc finger (ZF; blue residues). Dashed oval indicates the proposed location of the N-terminal antitermination region (ATR-N).

See also Figure S4.

Figures 4B and 4C) and the surface defined by the FIR are adjacent to each other, do not overlap each other, and are displayed on opposite faces of λQ (Figures 4B and 4C). Thus, the structure provides an explanation for the observation that λQ can simultaneously bind DNA (the QBE) and RNAP (the β flap tip) during λQ-loading (Deighan et al., 2008). In particular, assuming that the same λQ protomer mediates interaction with both the QBE and the β flap tip, λQ would be able to interact simultaneously with the QBE through the DIR and with the β flap tip through the FIR. Furthermore, the predicted distance between the QBE (bound to the DIR) and the β flap tip (bound to the FIR) would be ~20 Å. This arrangement is reminiscent of a transcription initiation complex, where the promoter −35 element and the β flap tip contact distinct regions of σR4, and the distance between the −35 element and the β flap tip is ~20 Å (Murakami et al., 2002). Thus, based upon the structure of λQ we propose that a tripartite QBE/β flap tip/λQ assembly forms during λQ-loading in a manner analogous to the tripartite promoter −35 element/β flap tip/σR4 assembly observed in a transcription initiation complex.

Mapping Residues Important for Interaction with the σ onto the Structure of λQ

During λQ-loading, λQ interacts with σR4 and stabilizes the binding of σR4 to the TTGACT motif. The TTGACT motif is located 1 bp upstream of the pause-inducing sequence, where σR2 is bound. Thus, the λQ-σR4 interaction stabilizes a nonstandard conformation of the transcription complex in which σR2 and σR4 are bound to promoter −−10-like and promoter −35-like DNA elements separated by a nonstandard-length spacer (1 bp spacer versus ~17 bp spacer; Figures 1B, 1C) (Devi et al., 2010; Nickels et al., 2002b). The interaction between λQ and σR4 stabilizes this nonstandard conformation of RNAP in which σR4 is displaced from its normal position on the tip of the β flap and is relocated, by ~50 to 60 Å to the base of the β flap (Devi et al., 2010).

The E134K substitution that increases the affinity of λQ for the QBE (Figure 3B) (Guo and Roberts, 2004) also disrupts the λQ-σR4 interaction (Figure S4B). Therefore, it appears likely that residue E134 is part of the region of λQ that interacts with σR4 (sigma interaction region, or SIR; magenta residue in Figures 4B and 4C). In the structure of λQ, residue E134 is located at the distal tip of a long, arm-like loop separated from the main mass of λQ by the zinc finger (Figures 4B and 4C). E134 is ~40 to 50 Å from the FIR and therefore is anticipated to be ~45 to 55 Å from the β flap tip when the β flap tip is bound to the FIR if no significant conformational change in λQ occurs upon DNA binding. The distance between residue 134 and the FIR (~40 to 50 Å) and the inferred distance between residue 134 and the β flap tip bound to the FIR (~45 to 55 Å) would be sufficient to enable a λQ protomer to bridge the distance between the β flap tip and σR4 bound to the TTGACT motif (~50 to 60 Å) and thereby to make simultaneous contact with the β flap tip and with σR4 bound to the TTGACT motif.

Mapping Residues Important for Antitermination onto the Structure of λQ

Guo and Roberts identified substitutions at eight residues within λQ (V6, G39, A50, C53, H56, L58, L202, and T206) that disrupt λQ's ability to antiterminate transcription (Deighan and Hochschild, 2007; Guo, 1999; Guo and Roberts, 2004). Two of these residues are present within the portion of λQ ordered in the

crystal structure and cluster near the C terminus (C-terminal antitermination region, or ATR-C; yellow residues in Figures 4B and 4C). ATR-C is located at the C terminus of the second α helix of the HTH motif. ATR-C is located at the opposite end of λQ from DIR/FIR because ATR-C and DIR/FIR are located at the C-terminal and N-terminal ends of this α helix, respectively, and because this α helix spans the length of λQ (compare yellow residues to red/green residues in Figures 4B and 4C). The most N-terminal residue of λQ ordered in the crystal structure (i.e., λQ residue 62) is relatively near to ATR-C (Figures 4B and 4C). We conjecture that the antitermination determinants specified by residues V6, G39, A50, C53, H56, and L58 within the N terminus of λQ (N-terminal antitermination region, or ATR-N; dashed circle in Figure 4B) are located adjacent to, and form a single antitermination determinant with, ATR-C.

Summary and Future Perspectives

λQ and the Q proteins of related lambdoid phages share a conserved mechanism of action and provide an excellent model system for studying mechanistic aspects of postinitiation gene control. Here we provide structural information for a Q antiterminator protein. The structure provides mechanistic insight into the loading step by which λQ joins the elongation complex. In particular, the structure suggests that during λQ-loading the same protomer of λQ can participate simultaneously in protein-DNA interaction with the QBE, protein-protein interaction with σR4, and protein-protein interaction with the β flap tip. The structure and characterization reported here set the stage for future experiments that will define the mechanisms by which λQ loads onto and alters the functional properties of the transcription elongation complex.

EXPERIMENTAL PROCEDURES

Protein Purification

$\lambda Q^{39–207}$ was purified following protocols developed by the Northeast Structural Genomics Consortium (Acton et al., 2011; Xiao et al., 2010) (see Supplemental Experimental Procedures).

Structure Determination and Refinement

Crystallization screening was performed using a microbatch-under-oil crystallization method at 4°C (Chayen et al., 1990). λQ crystals were grown in drops composed of 1.0 μl protein and 1.0 μl precipitant solution (100 mM ammonium bromide, 100 mM N-cyclohexyl-3-aminopropanesulfonic acid, 12% [weight per volume] polyethylene glycol 20000 [pH 10.0]) under paraffin oil. Crystals were cryoprotected with 15% ethylene glycol prior to flash-freezing in liquid nitrogen for data collection. A selenomethionyl SAD data set was collected at beamline X4A at the National Synchrotron Light Source at 100 K at the Se K-edge (λ = 0.97931 Å). Data were processed with the HKL2000 package (Otwinowski and Minor, 1997). The structure was solved by SAD using the program PHENIX (Adams et al., 2002). The model was completed using iterative cycles of manual rebuilding in Coot (Emsley and Cowtan, 2004) and refined against 2.1 Å data with PHENIX (Adams et al., 2002). The quality of the model was inspected by PROCHECK (Laskowski et al., 1993). The data processing and refinement statistics are summarized in Table 1. The atomic coordinates and structure factors have been deposited in the Protein Data Bank with the accession code 4MO1.

β Galactosidase Assays

Cells were transformed with the appropriate plasmids. Individual transformants were grown in lysogeny broth containing isopropyl-β-D-thiogalactoside (IPTG) and antibiotics at the indicated concentrations: carbenicillin (100 μg/ml), chloramphenicol (25 μg/ml), kanamycin (50 μg/ml), or tetracycline (10 μg/ml). Assays were done as described (Deighan et al., 2008) using microtiter plates and a microtiter plate reader. Graphs represent the values and SEM obtained from quadruplicate measurements.

Strains and Plasmids

A list of strains and plasmids is provided in Table S1.

ACCESSION NUMBERS

The Protein Data Bank accession number for the structure of $\lambda Q^{62–206}$ reported in this paper is 4MO1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.str.2013.12.010.

ACKNOWLEDGMENTS

We thank Jeff Roberts for λQ antibody and Irina Vvedenskaya for purified λQ^{FL}. This work is supported by National Institutes of Health (NIH) grants U54-GM094597 (to G.T.M. and L.T.), GM041376 (to R.H.E.), and GM088343 (to B.E.N.).

Received: November 8, 2013
Revised: December 18, 2013
Accepted: December 20, 2013
Published: January 16, 2014

REFERENCES

Acton, T.B., Xiao, R., Anderson, S., Aramini, J., Buchwald, W.A., Ciccosanti, C., Conover, K., Everett, J., Hamilton, K., Huang, Y.J., et al. (2011). Preparation of protein samples for NMR structure, function, and small-molecule screening studies. Methods Enzymol. *493*, 21–60.

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. *58*, 1948–1954.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA *98*, 10037–10041.

Chayen, N.E., Shaw Stewart, P.D., Maeder, D.L., and Blow, D.M. (1990). An automated system for micro-batch protein crystallization and screening. J. Appl. Cryst. *23*, 297–302.

Deighan, P., and Hochschild, A. (2007). The bacteriophage lambdaQ antiterminator protein regulates late gene expression as a stable component of the transcription elongation complex. Mol. Microbiol. *63*, 911–920.

Deighan, P., Diez, C.M., Leibman, M., Hochschild, A., and Nickels, B.E. (2008). The bacteriophage lambda Q antiterminator protein contacts the β-flap domain of RNA polymerase. Proc. Natl. Acad. Sci. USA *105*, 15305–15310.

Devi, P.G., Campbell, E.A., Darst, S.A., and Nickels, B.E. (2010). Utilization of variably spaced promoter-like elements by the bacterial RNA polymerase holoenzyme during early elongation. Mol. Microbiol. *75*, 607–622.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Goulet, A., Blangy, S., Redder, P., Prangishvili, D., Felisberto-Rodrigues, C., Forterre, P., Campanacci, V., and Cambillau, C. (2009). Acidianus filamentous virus 1 coat proteins display a helical fold spanning the filamentous archaeal viruses lineage. Proc. Natl. Acad. Sci. USA *106*, 21155–21160.

Guo, J. (1999). Study of lambda Q functional domains. PhD thesis (Ithaca, NY: Cornell University).

Guo, J., and Roberts, J.W. (2004). DNA binding regions of Q proteins of phages lambda and phi80. J. Bacteriol. *186*, 3599–3608.

Holm, L., and Sander, C. (1995). Dali: a network tool for protein structure comparison. Trends Biochem. Sci. *20*, 478–480.

Kuznedelov, K., Minakhin, L., Niedziela-Majka, A., Dove, S.L., Rogulja, D., Nickels, B.E., Hochschild, A., Heyduk, T., and Severinov, K. (2002). A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. Science *295*, 855–857.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. *26*, 283–291.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. Science *296*, 1285–1290.

Nickels, B.E., Dove, S.L., Murakami, K.S., Darst, S.A., and Hochschild, A. (2002a). Protein-protein and protein-DNA interactions of σ70 region 4 involved in transcription activation by lambda cI. J. Mol. Biol. *324*, 17–34.

Nickels, B.E., Roberts, C.W., Sun, H., Roberts, J.W., and Hochschild, A. (2002b). The σ^{70} subunit of RNA polymerase is contacted by the λQ antiterminator during early elongation. Mol. Cell *10*, 611–622.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Ring, B.Z., Yarnell, W.S., and Roberts, J.W. (1996). Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. Cell *86*, 485–493.

Roberts, J.W., Yarnell, W., Bartlett, E., Guo, J., Marr, M., Ko, D.C., Sun, H., and Roberts, C.W. (1998). Antitermination by bacteriophage lambda Q protein. Cold Spring Harb. Symp. Quant. Biol. *63*, 319–325.

Shankar, S., Hatoum, A., and Roberts, J.W. (2007). A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. Mol. Cell *27*, 914–927.

Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., and Artsimovitch, I. (2007). Structural basis for transcription elongation by bacterial RNA polymerase. Nature *448*, 157–162.

Xiao, R., Anderson, S., Aramini, J., Belote, R., Buchwald, W.A., Ciccosanti, C., Conover, K., Everett, J.K., Hamilton, K., Huang, Y.J., et al. (2010). The high-throughput protein sample production platform of the Northeast Structural Genomics Consortium. J. Struct. Biol. *172*, 21–33.

Yarnell, W.S., and Roberts, J.W. (1999). Mechanism of intrinsic transcription termination and antitermination. Science *284*, 611–615.

Young, B.A., Anthony, L.C., Gruber, T.M., Arthur, T.M., Heyduk, E., Lu, C.Z., Sharp, M.M., Heyduk, T., Burgess, R.R., and Gross, C.A. (2001). A coiled-coil from the RNA polymerase β' subunit allosterically induces selective nontemplate strand binding by σ^{70}. Cell *105*, 935–944.